

IN THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application:

Claim 1 (currently amended): A document processor which displays and prints in a predetermined format a plurality of document data input thereto, comprising:

a document memory which stores input document data;

a selection unit which selects allows a user to select all or part of document data

stored in said documents document memory;

a characteristics extraction unit which extracts data relating to characteristics of letter rows from all or part of the document data selected by said selection unit;

a work processing unit which work-processes all or part of the document data based on the data relating to characteristics of letter rows extracted by said characteristics extraction unit; and

an output unit which outputs all or part of the document data work-processed by said work processing unit,

wherein processing performed by said work processing unit include classifying the characteristics, retrieving the characteristics, rearranging the characteristics, extracting a representative value, a maximum value and a minimum value from the characteristics, and calculating the characteristics.

Claim 2 (currently amended): The document processor according to claim 1, wherein said output unit comprises an item value set unit which sets a plurality of item values based on the contents of all or part of the document data work-processed by said work-processing

Application No. 09/472,249
Reply to Office Action of August 13, 2003

unit[[;]] and a totalization unit which totalizes all or part of the document data for each of the plurality of item [[value]] values set by said item value set unit[[;]], said output unit outputs outputting all or part of the document data in the format of a table having an item value as at least one axis.

Claim 3 (original): The document processor according to claim 1, wherein said output unit outputs all or part of the document data work-processed by said work-processing unit together with all or part of the document data in its state prior to work-processing by said work processing unit.

Claim 4 (original): The document processor according to claim 1, wherein said document memory further stores all or part of the document data work-processed by said work processing unit.

Claim 5 (original): The document processor according to claim 1, wherein said selection unit further selects all or part of the document data output by said output unit.

Claim 6 (currently amended): The document processor according to claim 1, wherein said document memory further stores data relating to contents of the work processing performed by said processing unit.

Claim 7 (canceled)

Claim 8 (currently amended): A document classification device which classifies documents based on contents thereof comprising:

an input unit which inputs a document data;

a language analyzer unit which analyzes document data input by said input unit and obtains language analysis information;

a vector creation unit which creates document characteristic vectors for the document data based on the language analysis information obtained by said language analyzer unit;

a classification unit which classifies documents based on the degree of similarity between document characteristic vectors created by said vector creation unit, and creates clusters of documents;

a cluster characteristics calculation unit which calculates cluster characteristics, which are characteristics of clusters of documents created by said classification unit;

a display unit which displays the cluster characteristics calculated by said cluster characteristics calculation unit;

a cluster selection specification unit which selects predetermined clusters from cluster of documents created by said classification unit; [[and]]

a classification category memory which stores cluster characteristics, calculated by said cluster characteristics calculation unit, as constituent elements of classification categories;

a document characteristic vector memory which stores document characteristic vectors created by said vector creation unit; and

a vector correction unit which corrects document characteristic vectors stored in said document characteristic vector memory so that document characteristic vectors of documents belonging to clusters selected by said cluster selection unit are deleted.

Claim 9 (currently amended): The document classification device according to claim 8, ~~further comprising document characteristic vector memory which stores document characteristic vectors created by vector creation unit; and vector correction unit which corrects document characteristic vectors stored in said document characteristic vector memory, so that document characteristic vectors of documents belonging to clusters selected by said cluster selection unit are deleted;~~ wherein said classification unit [[which]] classifies documents based on the document characteristic vectors corrected by said vector correction unit.



Claim 10 (currently amended): The document classification device according to claim 8, ~~further comprising document characteristic vector memory which stores document characteristic vectors created by vector creation unit; and~~ a document expression space correction unit which corrects document expression space when determining the degree of similarity between document characteristic vectors stored in said document characteristic vectors memory, based on a characteristics amount calculated from clusters selected by said cluster selection unit[[;]], wherein said classification unit ~~classify~~ classifies the documents based on the degree of similarity between document characteristic vectors created by said vector creation unit, using the document expression space corrected by said document expression space correction unit.

Claim 11 (canceled)

Claim 12 (currently amended): The document classification device according to claim 8, further comprising a selection information appending unit which appends selection information showing the fact of selection when all or part of the documents belonging to a cluster of documents created by said classification unit have been selected[[;]],

wherein said display unit displays the cluster characteristics, and ~~also displays~~ the selection information appended by said selection information appending unit.

Claim 13 (original): The document classification device according to claim 8, wherein said classification category memory stores cluster characteristics and/or information created by an operator, in addition to all or part of the documents belonging to a cluster of documents selected by said selection specification unit, as constituent elements of classification categories.

Claim 14 (currently amended): A document classification device which classifies document clusters in accordance with contents thereof comprising:

a document input unit which inputs document data groups;
a document dividing unit which divides document data into one or multiple divided document data based on a predetermined reference;
a document-divided document map creation unit which creates a map showing the correspondence between the document data and the divided document data;
a divided document classification unit which classifies the divided document data such that multiple topics and meanings in a document data are classified into categories according to specific topics and meaning;

a divided document classification result creation unit which creates divided document classification result information based on a classification result of said divided document classification unit; and

a document classification result creation unit which creates classification result information of the [[above]] document data using the document-divided document map and the divided document classification result information.

Claim 15 (currently amended): The document classification device according to claim 14, further comprising a document save unit which saves the document data[[:]], a divided document save unit which saves the divided document data[[:]] and a document-divided document map save unit which saves a document-divided document map created by said document-divided document map creation unit.

Claim 16 (original): The document classification device according to claim 15, further comprising divided document classification result save unit which saves the divided document classification result information created by said divided document classification result creation unit.

Claim 17 (original): The document classification device according to claim 14, wherein a plurality of divided document data created by said document dividing unit comprises the document data in its state prior to being divided.

Claim 18 (original): The document classification device according to claim 14, wherein said document dividing unit divides document data based on information relating to the structure of the document data.

Claim 19 (currently amended): The document classification device according to claim 14, further comprising a document element extraction unit which extracts elements in the document data[[;]] and an element-accompanying information extraction unit which extracts element-accompanying information accompanying the elements extracted by said document element extraction unit[;]], said document dividing unit divides dividing the document data using elements extracted by said document element extraction unit, or the elements and element-accompanying information extracted by said element-accompanying information extraction unit.

Claim 20 (original): The document classification device according to claim 14, wherein said document dividing unit divides the document data in compliance with a specified specification range.

Claim 21 (original): The document classification device according to claim 14, wherein said document dividing unit divides the document data based on the number of letters, the number of sentences, or both the number of letters and the number of sentences.

Claim 22 (original): The document classification device according to claim 14, wherein said document classification result creation unit extracts and presents information

Application No. 09/472,249
Reply to Office Action of August 13, 2003

showing document data, and representative information accompanying the document data, as classification result information.

Claim 23 (original): The document classification device according to claim 22, wherein said document classification result creation unit extracts and presents information showing divided document data, and representative information accompanying the divided document data, as classification result information.

Claim 24 (currently amended): A document processing method which outputs a plurality of input document data in order to display or print the document data in a predetermined format, comprising the steps of:

storing input document data;
selecting all or part of the document data stored in the storing step;
extracting data relating to characteristics of letter rows from all or part of the document data selected in the selection step;
work-processing all or part of the document data based on the data relating to characteristics of letter rows extracted in the characteristics extraction step; and
outputting all or part of the document data work-processed in the work-processing step,

wherein processing performed by said work processing unit include classifying the characteristics, retrieving the characteristics, rearranging the characteristics, extracting a representative value, a maximum value and a minimum value from the characteristics, and calculating the characteristics.

Claim 25 (currently amended): The document processing method according to claim 24, wherein the step of outputting comprises the steps of setting a plurality of item values based on the contents of all or part of the document data work-processed in the work-processing step[[; and]], totalizing all or part of the document data for each of the plurality of item [[value]] values set by in the item value setting step[[;]], and
outputs outputting all or part of the document data in the format of a table having an item value as at least one axis.

Claim 26 (original): The document processing method according to claim 24, wherein the step of outputting comprises outputting all or part of the document data work-processed in the work-processing step together with all or part of the document data in its state prior to work-processing in the work-processing step.

Claim 27 (original): The document processing method according to claim 24, wherein the step of storing further comprises storing all or part of the document data work-processed in the work-processing step.

Claim 28 (original): The document processing method according to claim 24, wherein the step of selecting further comprises selecting all or part of the document data output in the output step.

Claim 29 (currently amended): The document processing method according to claim 24, wherein the step of storing a document further comprises storing data relating to contents of the work processing performed by said processing step.

Claim 30 (canceled)

Claim 31 (currently amended): A document classification method of classifying documents based on contents thereof, comprising the steps of:

inputting a document data;

language-analyzing document data input in the step of inputting and obtaining language analysis information;

creating document characteristic vectors for the document data based on the language analysis information obtained in the step of language-analyzing;

classifying documents based on the degree of similarity between document characteristic vectors created in the step of creating vectors, and creating clusters of documents;

calculating cluster characteristics, which are characteristics of clusters of documents created in the step of classifying;

displaying the cluster characteristics calculated in the step of calculating cluster characteristics;

selecting predetermined clusters from cluster of documents created in the step of classifying; [[and]]

storing cluster characteristics, calculated in the step of calculating cluster characteristics, as constituent elements of classification categories;

storing document characteristics vector created in the step of creating document characteristic vectors; and

correcting document characteristic vectors stored in the step of storing document characteristic vectors, so that document characteristic vectors of documents belonging to clusters selected by the step of selecting clusters are deleted.

Claim 32 (currently amended): The document classification method according to claim 31, ~~further comprising the step of correcting document characteristic vectors stored in the step of storing document characteristic vectors, so that document characteristic vectors of documents belonging to clusters selected by the step of selecting clusters are deleted;~~

wherein the step of classifying comprising classifying documents based on the document characteristic vectors corrected by the step of correcting vectors.

Claim 33 (currently amended): The document classification method according to claim 31, further comprising a step of correcting document expression space when determining the degree of similarity between document characteristic vectors stored in the step of storing document characteristic vectors, based on a characteristics amount calculated from clusters selected in the step of selecting clusters[[;]],

wherein the step of classifying ~~comprising~~ comprises classifying documents based on the degree of similarity between document characteristic vectors created in the step of creating vectors, using the document expression space corrected in the step of correcting the document expression space.

Claim 34 (canceled)

Claim 35 (currently amended): The document classification method according to claim 31, further comprising the steps of appending selection information showing the fact of selection when all or part of the documents belonging to a cluster of documents created in the step of classifying have been selected[[:]],

wherein the step of displaying comprising displaying the cluster characteristics, and also displaying the selection information appended in the step of appending selection information.

Claim 36 (original): The document classification device according to claim 31, wherein the step of creating classification categories comprises creating cluster characteristics and/or information created by an operator, in addition to all or part of the documents belonging to a cluster of documents selected in the step of specifying selection, as constituent elements of classification categories.

Claim 37 (currently amended): A document classification method which classifies document clusters in accordance with contents thereof comprising the steps of:
inputting document data groups;
dividing document data into one or multiple divided document data based on a predetermined reference;
creating a map showing the correspondence between the document data and the divided document data;
classifying the divided document data such that multiple topics and meanings in document data are classified into categories according to specific topics and meanings;

creating divided document classification result information based on the classification result of classifying the divided documents; and
creating classification result information of the document data using the document-divided document map and the divided document classification result information.

Claim 38 (currently amended): A computer-readable recording medium in which is stored programs for executing a document classification method, which document classification method comprising the steps of:

storing input document data;
selecting all or part of the document data stored in the storing step;
extracting data relating to characteristics of letter rows from all or part of the document data selected in the selection step;
work-processing all or part of the document data based on the data relating to characteristics of letter rows extracted in the characteristics extraction step; and
outputting all or part of the document data work-processed in the work-processing step,

wherein processing performed by said work processing unit include classifying the characteristics, retrieving the characteristics, rearranging the characteristics, extracting a representative value, a maximum value and a minimum value from the characteristics, and calculating the characteristics.

Claim 39 (canceled)

Claim 40 (currently amended): A computer-readable recording medium in which is stored programs for executing a document classification method, which document classification method comprising the steps of:

inputting a document data;

language-analyzing document data input in the step of inputting and obtaining language analysis information;

creating document characteristic vectors for the document data based on the language analysis information obtained in the step of language-analyzing;

classifying documents based on the degree of similarity between document characteristic vectors created in the step of creating vectors, and creating clusters of documents;

calculating cluster characteristics, which are characteristics of clusters of documents created in the step of classifying;

displaying the cluster characteristics calculated in the step of calculating cluster characteristics;

selecting predetermined clusters from cluster of documents created in the step of classifying; [[and]]

storing cluster characteristics, calculated in the step of calculating cluster characteristics, as constituent elements of classification categories;

storing document characteristics vector created in the step of creating document characteristic vectors; and

correcting document characteristic vectors stored in the step of storing document characteristic vectors, so that document characteristic vectors of documents belonging to clusters selected by the step of selecting clusters are deleted.

Claim 41 (currently amended): A computer-readable recording medium in which is stored programs for executing a document classification method, which document classification method comprising the steps of:

inputting document data groups;

dividing document data into one or multiple divided document data based on a predetermined reference;

creating a map showing the correspondence between the document data and the divided document data;

classifying the divided document data such that multiple topics and meanings in document data are classified into categories according to specific topics and meanings;

creating divided document classification result information based on the classification result of classifying the divided documents; and

creating classification result information of the document data using the document-divided document map and the divided document classification result information.